OPEN ACCESS

University of BRISTOL

Anantrasirichai, N., Agrafiotis, D., & Bull, DR. (2008). A concealment based approach to distributed video coding. In *15th IEEE International Conference on Image Processing, 2008 (ICIP 2008), San Diego* (pp. 2232 - 2235). Institute of Electrical and Electronics Engineers (IEEE). https://doi.org/10.1109/ICIP.2008.4712234

Peer reviewed version

Link to published version (if available):
10.1109/ICIP.2008.4712234

Link to publication record in Explore Bristol Research
PDF-document

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/

# A Concealment based Approach to Distributed Video Coding


*Nantheera Anantrasirichai,*
*Dimitris Agrafiotis,*
*Dave Bull*

Signal Processing Group
Centre for Communications Research
Dept.of Electrical & Electronic Engineering

# Contents of Presentation

- Introduction
  - ⇨ Distributed Video Coding (DVC)
- Proposed schemes
  - ⇨ Hybrid Key/WZ frames
  - ⇨ Block-based concealment
- Results
- Conclusions

# Introduction

- **Distributed Video Coding (DVC)**
  - Very low complex video transmitters
  - Two information theory principles: Slepian-Wolf (SW) and Wyner-Ziv (WZ)
  - Separate encoding, but joint decoding
  - Single source video frames are split into two categories, key frames and WZ frames

# Introduction

- **Distributed Video Coding (DVC)**
  - Key frames are intra coded with a conventional encoder and are made available at the decoder.
  - WZ frame coding, which may involve transformation, uses quantisation followed by channel coding applied.
  - At the decoder, the key frames are used for creating an estimate of the WZ frames (side information)
  - The side information is seen as a noisy version of the original coded WZ frame.
  - Finally, the received parity bits are used to correct the errors present in this noisy version of the WZ data.

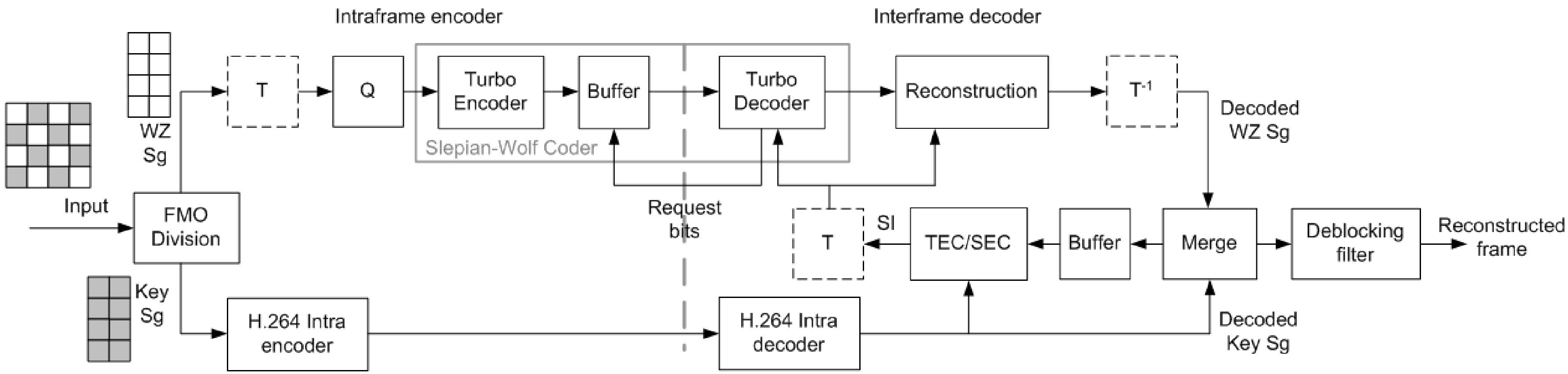
# Proposed System

This paper presents a concealment based approach to distributed video coding with…

- hybrid Key/WZ frames via an FMO type interleaving of macroblocks.
- spatio-temporal concealment for generating the side information on a MB basis
- mode selection for switching between the two concealment approaches and for deciding how the correlation noise is estimated
- local (MB wise) correlation noise estimation
- modified B frame quantisation

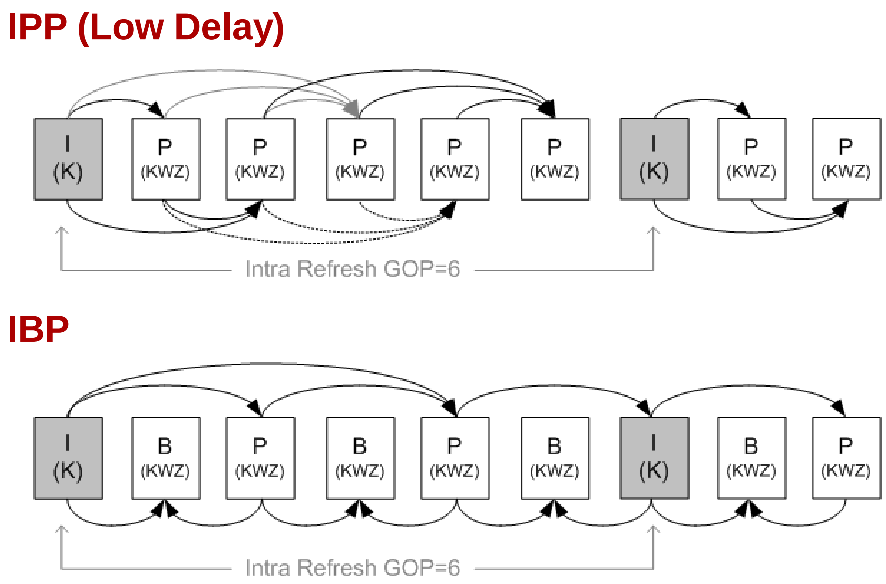
# Proposed System

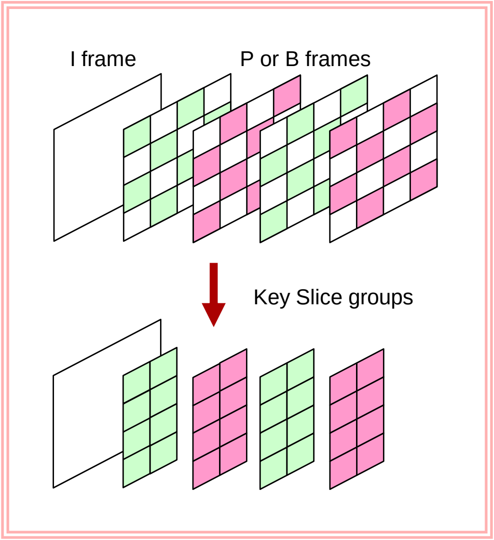
- FMO DVC codec



# Proposed System

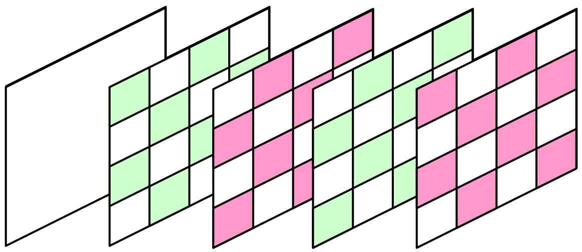
- GOP Structure

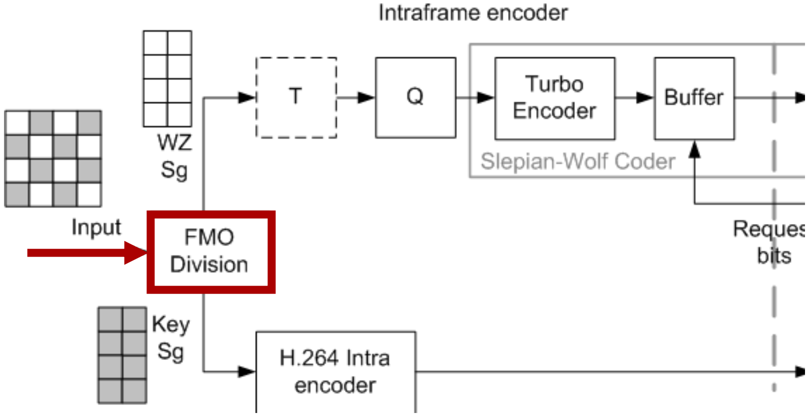


# Proposed System Encoder

- Interleaving

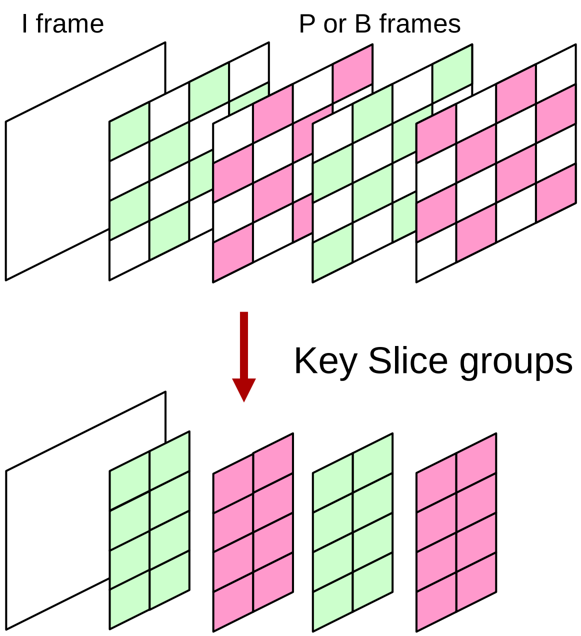


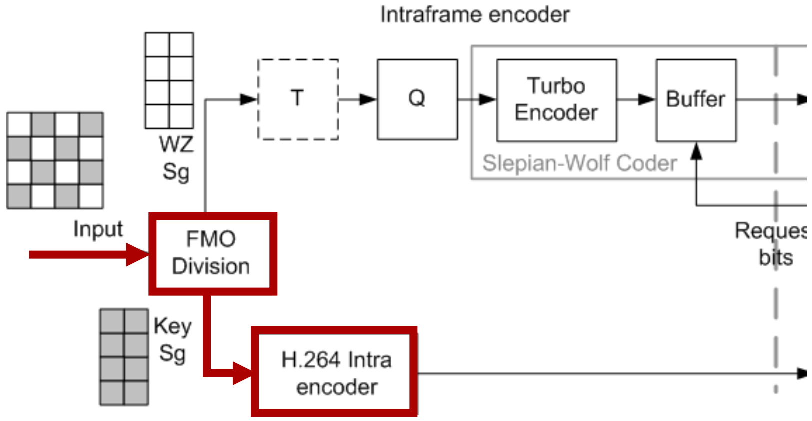




The first step involves splitting of the current input frame into KEY and WZ groups, in a similar fashion to the dispersed type of flexible macroblock ordering (FMO) specified in H.264.

# Proposed System Encoder

- KEY group coding

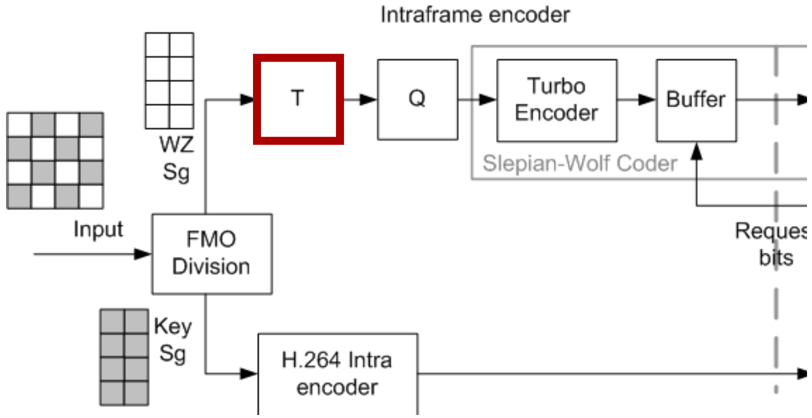



Then, the KEY MBs are, horizontally shifted to make a new frame of the same height but half the width of the original which is then encoded with H.264 in intra mode.

# Proposed System Encoder

- WZ group coding




***Transform*** might be applied to remove spatial redundancy at the expense of a slight increase in complexity.

In addition to the matrices defined in [1] and [2], two more matrices (shown below) are proposed for very low bitrates.

| | | | |
|---|---|---|---|
| 8 | 2 | 0 | 0 |
| 2 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 |

| | | | |
|---|---|---|---|
| 12 | 4 | 0 | 0 |
| 4 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 |

Note than the use of 0 quantisation levels implies that no bits are sent.

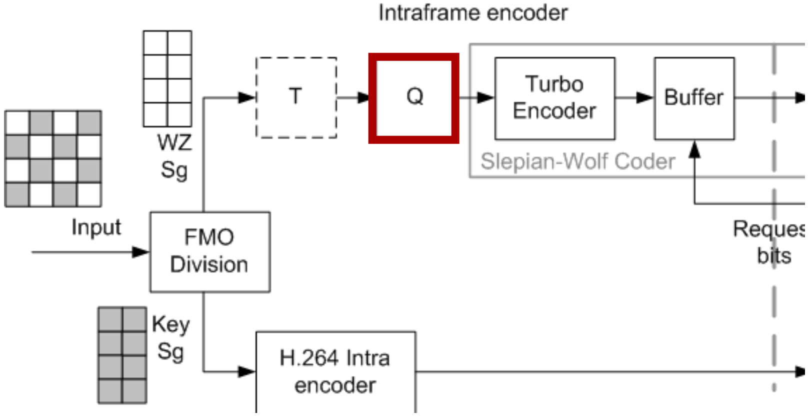

[1] A. Aaron, S. Rane, E. Setton, and B. Girod, “Transform-domain Wyner–Ziv codec for video,” presented at the SPIE Visual Communications and Image Processing Conf., San Jose, CA,2004.

[2] C. Brites, J. Ascenso, and F. Pereira, “Improving Transform Domain Wyner-Ziv Video Coding Performance,” in *proc IEEE ICASSP*, 2006, pp. II-525-528.

# Proposed System Encoder

- B-frame quantisation

To further reduce this bitrate we increase the quantisation step size of these WZ MBs relative to the step size used for the WZ MBs of the P frames.




The relationship between the number of quantisation levels $Q_B^n$ and $Q_P^n$ for B and P frames respectively, for *n* bit-planes can be written as

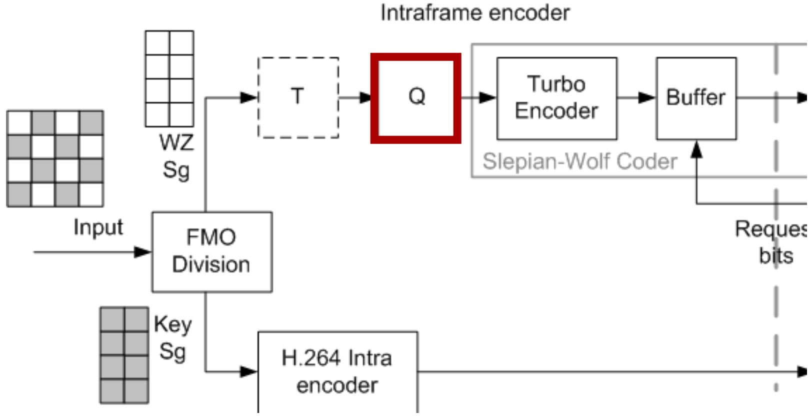
$$Q_B^n = \left| a \cdot Q_P^n \right|$$

Based on experimental results, we have selected *a* so that $Q_B^n = 1/2(Q_P^n + Q_P^{n-1})$ i.e. *a* = 0.75.

# Proposed System Encoder

- B-frame quantisation

| **# of bit-planes (*n*)** | $Q_P^n$ | $Q_B^n$ |
|---|---|---|
| 4 | 16 | 12 |
| 3 | 8 | 6 |
| 2 | 4 | 3 |
| 1 | 2 | 2 |

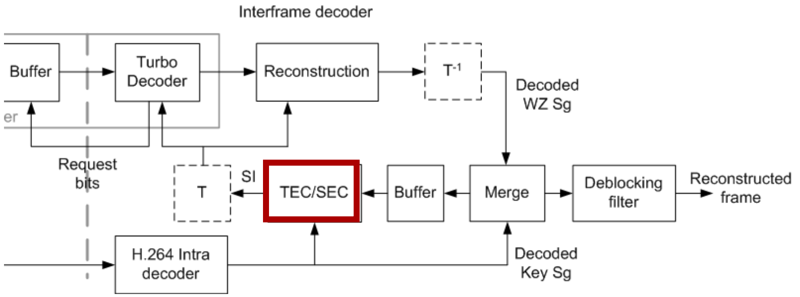



Bits spent for each decoded bit-plane of a B-Frame WZ group for different number of quantisation levels (BER $10^{-3}$)

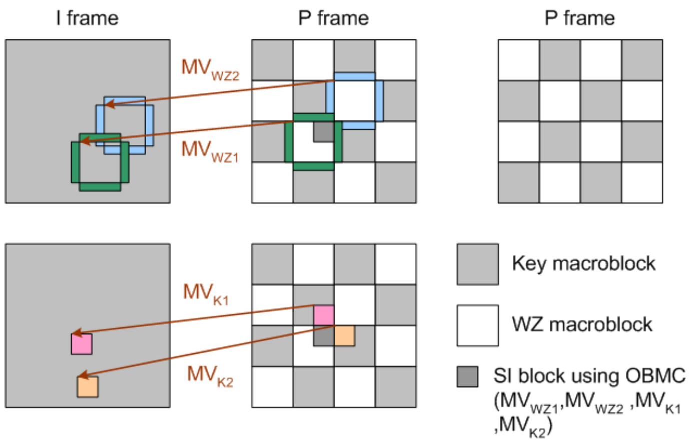
| **Bit-plane** | $Q_B^3 = Q_P^3$ | $Q_B^3 = 0.75 Q_P^3$ |
|---|---|---|
| 1 (MSB) | 12672 | 6336 |
| 2 | 22176 | 12672 |
| 3 (LSB) | 31680 | 31680 |

# Proposed System Decoder

- Side Information Generation



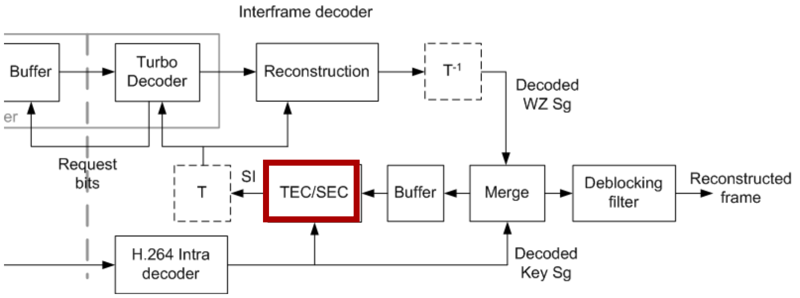

**Temporal Error Concealment**



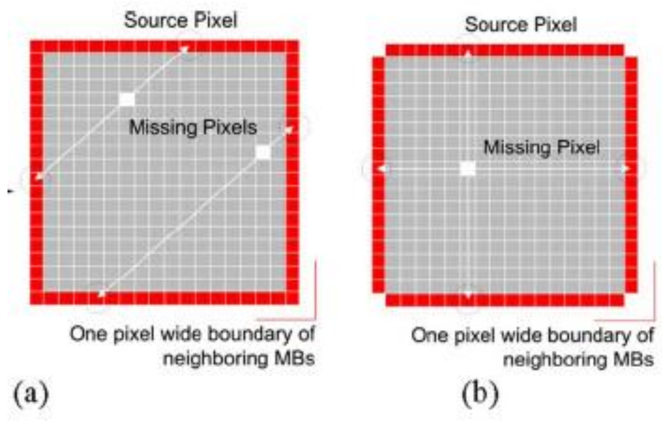

Each missing WZ block is divided into four sub-blocks. Each sub-block is concealed using two motion vectors of neighbouring KEY blocks and two motion vectors generated using external boundary matching error (EBME) of its own WZ block and adjacent WZ blocks. These four possible sub-blocks are blended using a cosine weighting matrix.

# Proposed System Decoder

- Side Information Generation




**Spatial Error Concealment**

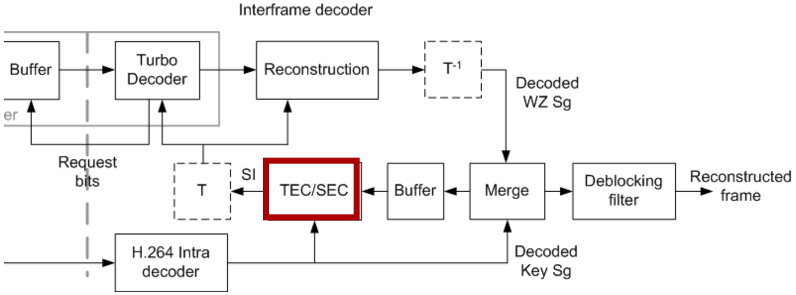



SEC uses bordering KEY pixels to conceal the missing WZ pixels of each WZ block through
(a) directional interpolation or
(b) bilinear interpolation
along detected edges.

D. Agrafiotis, D. R. Bull, N. Canagarajah, “Spatial error concealment with edge related perceptual considerations,” Signal Processing: image communication, vol.21, iss.2, Feb 2006, pp 130-142.

# Proposed System Decoder

- Side Information Generation




**Mode Selection**

- Evaluating the levels of motion compensated activity (TA) and spatial activity (SA) in the neighbourhood of that block

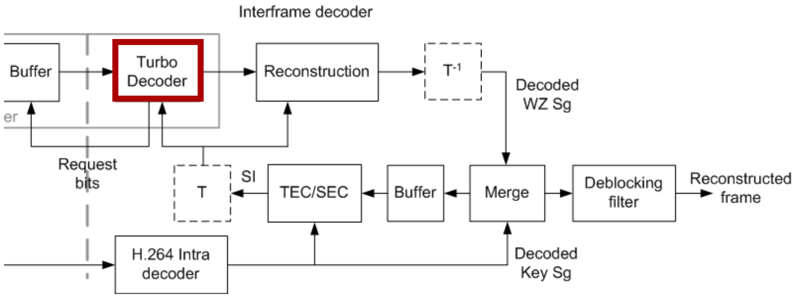
  $SA = E[(x - \mu)^2]$ and $TA = E[(x - x^*)^2]$

  If (SA<TA) and (TA>3), MB is processed by SEC.

TA = mean squared error between the key blocks surrounding the missing block in the current frame and those surrounding the replacement block in the reference frame

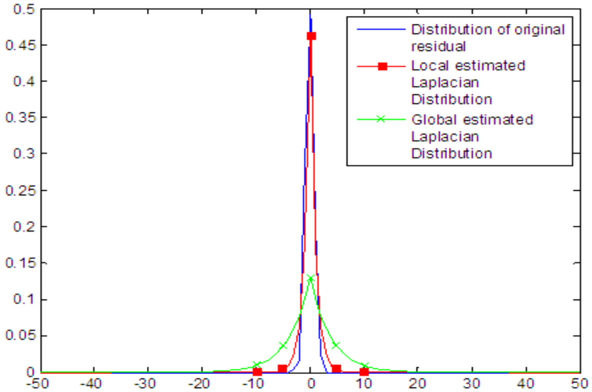
SA = variance of the surrounding key blocks in the current frame

# Proposed System Decoder




- Correlation Noise Estimation

Using the 4-neighbouring Key macroblocks of each WZ MB.



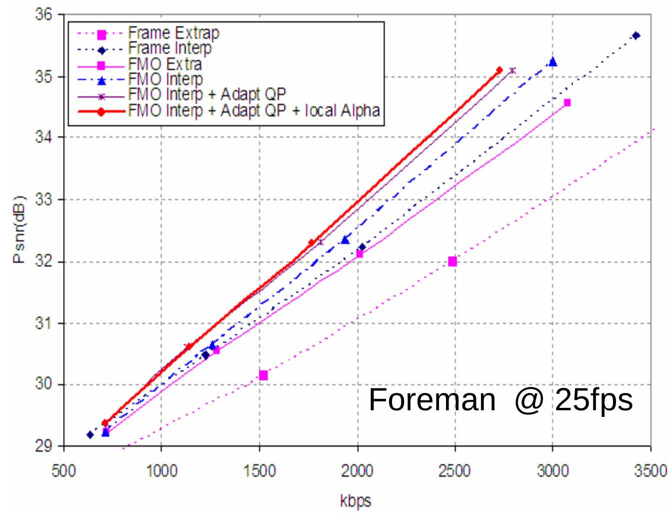

**Correlation noise estimation example**

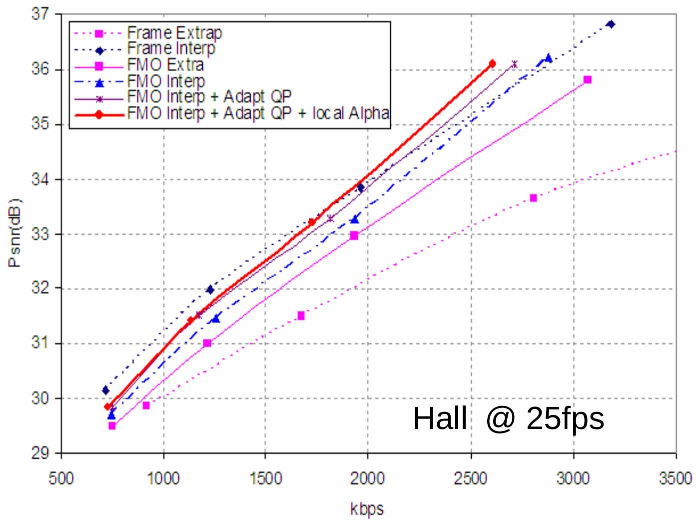
**Average bits spending on the first and second bitplanes when local or global noise distribution is used**

| Sequence | Global | | Local | | % improvement | |
|---|---|---|---|---|---|---|
| | 1st | 2nd | 1st | 2nd | 1st | 2nd |
| Akiyo | 8673 | 15684 | 8413 | 13191 | 3.0% | 15.9% |
| Foreman | 8933 | 15840 | 6855 | 12776 | 23.2% | 19.1% |
| Breakdancers | 6024 | 18696 | 3428 | 16100 | 43.1% | 13.9% |

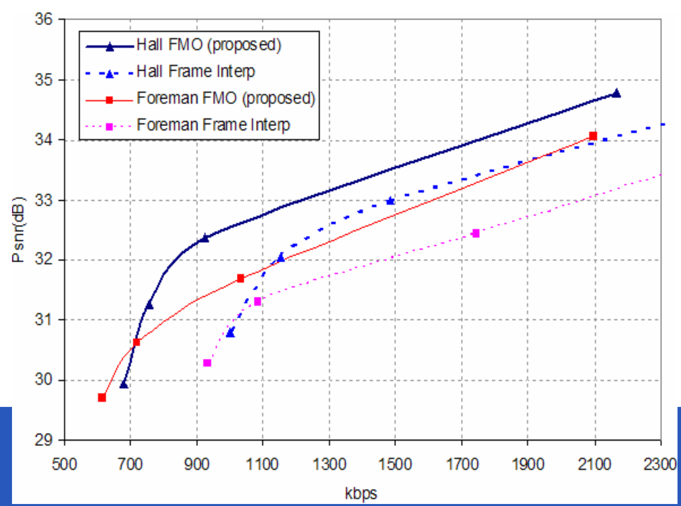
# Results

Pixel domain DVC performance



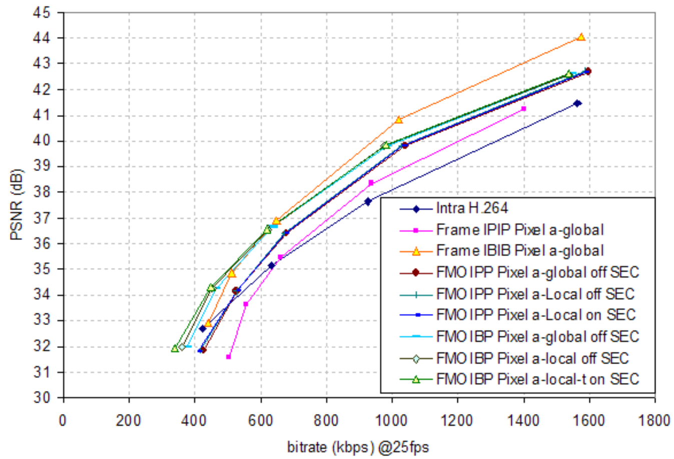



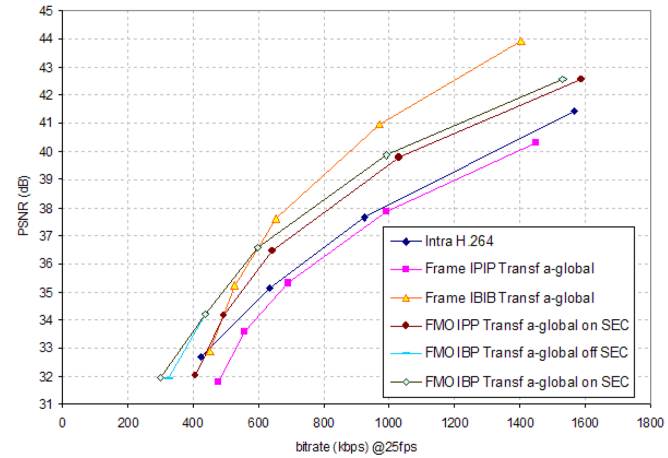

Transform domain DVC performance



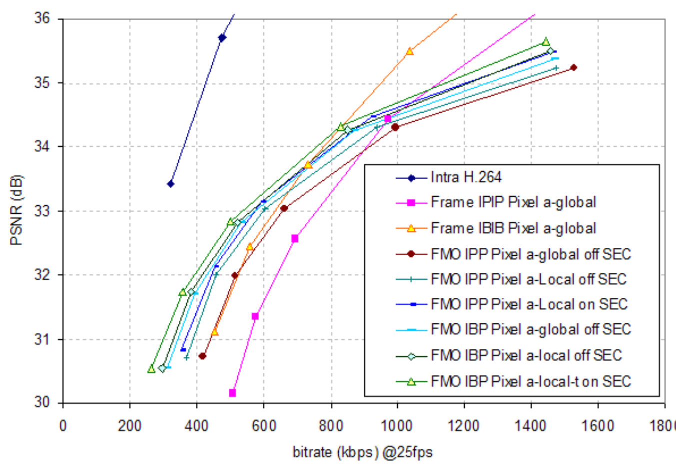
# Results

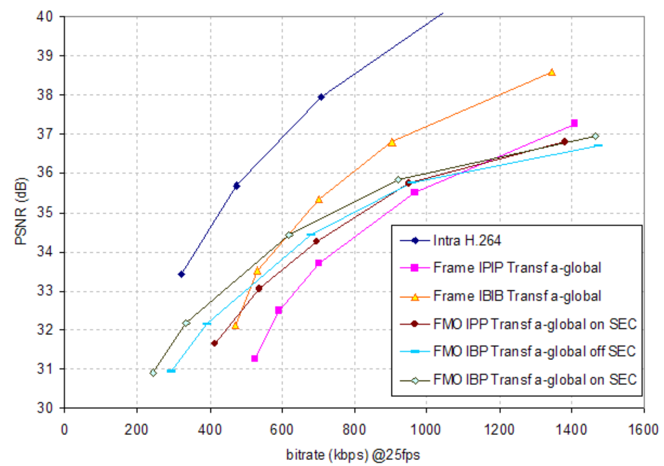
Rate-distortion performance of *Akiyo*



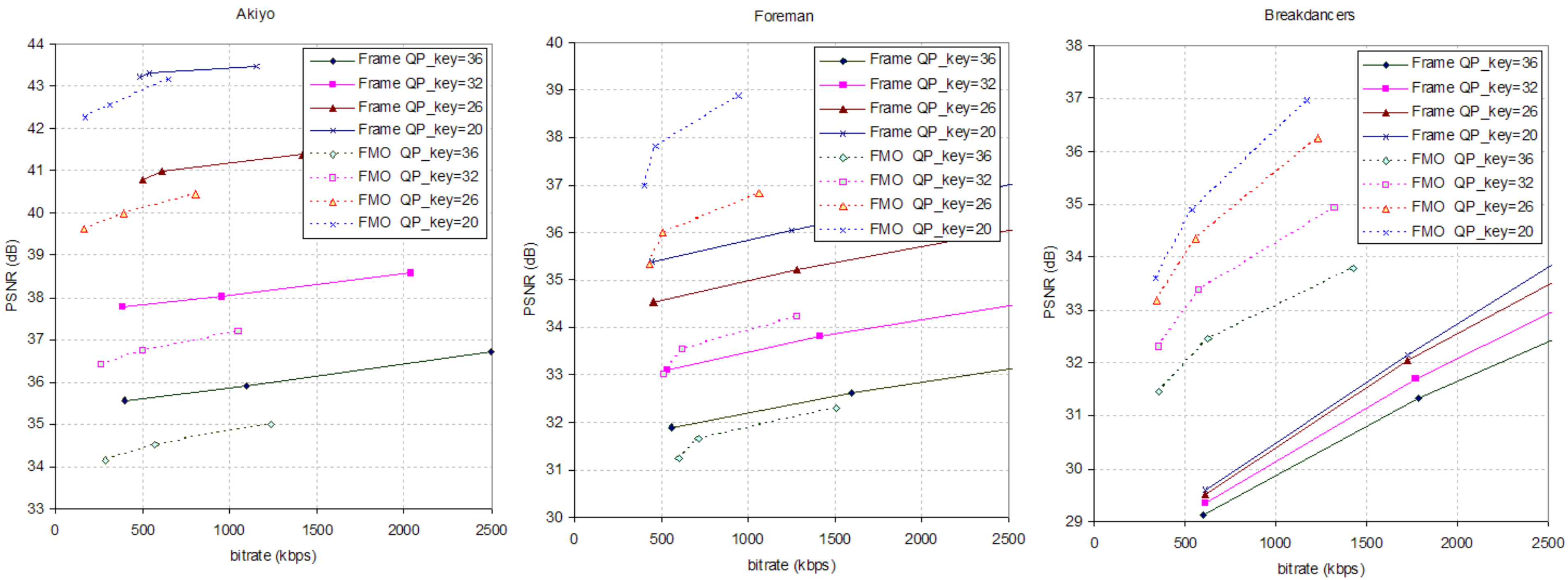

Rate-distortion performance of *Breakdancers*



# Results

WZ performance



# Conclusions

## This paper presents

- Extension of the Key-Wyner/Ziv framework to the IBP and transform domain DVC scenario
- Use of spatio-temporal concealment for generating the side information on a MB basis
- Mode selection for switching between the two concealment approaches and for deciding how the correlation noise is estimated
- Local (MB wise) correlation noise estimation
- Modified B frame quantisation

# Thank you